Comparative activity and functional ecology of permafrost soils and lithic niches in a hyper-arid polar desert

Authors: J. Goordial[1*], A. Davila[2], C.W Greer [3], R. Cannam[1], J. DiRuggiero[4], C.P. McKay[2] and L.G. Whyte[1]

Affiliations:
[1] McGill University Macdonald Campus, Ste-Anne-de-Bellevue, Quebec, Canada.
[2] NASA Ames Research Center, Moffett Field, CA, USA.
[3] National Research Council Canada, Montreal, QC, Canada
4 John Hopkins University, Baltimore, MD, USA

*Corresponding author: Jacqueline.goordial@mail.mcgill.ca
Department of Natural Resource Sciences, McGill University, Macdonald Campus, 21,111 Lakeshore, Ste-Anne-de-Bellevue, Quebec, Canada H9X 3V9.
Tel: 514-398-7881 Fax: 514-398-7990

Running Head: Ecology of Dry Valley permafrost and cryptoendoliths:

Originality-Significance Statement: This study represents the first metagenomic interrogation of Antarctic permafrost and polar cryptoendolithic microbial communities. The results underlie two different habitability conditions in the same location under extreme cold and dryness: the permafrost habitat where viable microbial life and activity is questionable, and the cryptoendolithic habitat which contains organisms capable of growth under the extreme conditions of the Antarctic Dry Valleys.

Summary:

Permafrost in the high elevation McMurdo Dry Valleys of Antarctica ranks amongst the driest and coldest on Earth. Permafrost soils appear to be largely inhospitable to active microbial life, but sandstone lithic microhabitats contain a trophically simple but functional cryptoendolithic community. We used metagenomic sequencing and activity assays to examine the functional capacity of permafrost soils and cryptoendolithic communities in University Valley, one of the most extreme regions in the Dry Valleys. We found metagenomic evidence that cryptoendolithic microorganisms are adapted to the harsh environment and capable of metabolic activity at *in situ* temperatures, possessing a suite of stress response and nutrient cycling genes to fix carbon under the fluctuating conditions that the sandstone rock would experience during the summer months. We additionally identified genes involved in microbial competition and cooperation within the cryptoendolithic habitat. In contrast, permafrost soils have a lower richness of stress response genes, and instead the metagenome is enriched in genes involved with dormancy and

sporulation. The permafrost soils also have a large presence of phage genes and genes involved in the recycling of cellular material. Our results underlie two different habitability conditions under extreme cold and dryness: the permafrost soils which is enriched in traits which emphasize survival and dormancy, rather than growth and activity; and the cryptoendolithic environment that selects for organisms capable of growth under extremely oligotrophic, arid, and cold conditions. This study represents the first metagenomic interrogation of Antarctic permafrost and polar cryptoendolithic microbial communities.

Introduction

A large fraction of Earth's biosphere is permanently cold, and cold adapted microorganisms capable of growth at temperatures well below freezing have been found in Polar and non-Polar Regions (De Maayer et al 2014). It is now well established that permafrost—ground that remains at or below 0°C for at least two consecutive years—can host viable and active communities of microorganisms (Bakermans et al 2014, Goordial et al 2013, Hultman et al 2015, Mackelprang et al 2011, Rivkina et al 2000, Steven et al 2008, Tuorto et al 2014), in addition to cells capable of resuming metabolic activity upon thawing, even after years of cryobiosis (Legendre et al 2014).

Current knowledge of the microbial diversity, metabolic activity, and ecology in permafrost is primarily informed by investigations of Arctic and Alpine regions, despite the fact that 37% of the world's permafrost exists in the Antarctic, and some of the coldest permafrost soils are found in the McMurdo Dry Valleys (Campbell and Claridge 1987, Marchant and Head 2007). Here, the extremely cold and dry environment results in a layer of dry permafrost soils overlaying ice-cemented soils (Campbell and Claridge, 1987) a condition that is rare on Earth. Microbiology investigations of dry surface soils throughout the Dry Valleys indicate the presence of localized, diverse edaphic bacterial communities (Aislabie et al 2006, Cary et al 2010, Chan et al 2013, Lee

et al 2012, Niederberger et al 2012, Pointing et al 2009, Smith et al 2006, Stomeo et al 2012). However, the abundance and diversity of these communities is strongly influenced by climate, and biomass in the inland, high elevation Dry Valleys, which are drier and colder, is significantly lower (10^3-10^4 cells g^{-1}) (Gilichinsky et al 2007, Goordial et al 2016) relative to maritime influenced Dry Valleys which are relatively wetter and warmer, and can harbour as much as 10^8 cells g^{-1} (Cowan et al 2002). Limited work has been carried out in ice-bearing permafrost soils (Bakermans et al 2014, Gilichinsky et al 2007, Goordial and Whyte 2014, Goordial et al 2016, Tamppari et al 2012) compared with dry surface soils, largely due to the logistical challenges associated with sampling ice-cemented ground in these very remote regions.

Some of the coldest and driest permafrost soils studied to date in the Dry Valleys are found in University Valley, a high elevation (1650-1800 m.a.sl) glacial valley in the Quartermain Range (Goordial et al 2016, Tamppari et al 2012). Air temperature data collected between 2009-2013 in University Valley was measured to be always below 0ºC, with a maximum, minimum and mean hourly air temperature of -2.8ºC, −45.5 ºC, and −23.4 ºC respectively (Goordial et al 2016). During the summer months (Dec - Feb) mean air temperature was -13.9ºC, with daily temperature fluctuations between -15ºC and -5ºC, depending on cloud cover and shadowing (Lacelle et al 2015). Permafrost soils in University Valley contain negligible microbial biomass (10^3 cells g^{-1}) and culturable organisms ($0-10^1$ CFU g^{-1}), and microbial activity in some of these soils can be undetectable *in situ* and in long-term microcosm assays (Goordial et al 2016). These soils are potentially devoid of any active microbial life, or alternatively, any existing metabolically active cells are below the detection limits of current methodologies.

Immediately adjacent to these depauperate soils there is a trophically simple but functional cryptoendolithic microbial community that occupies the pore space of sandstone rocks and cliffs, and which is comprised of photoautotrophs, lichenizing and free-living fungi and heterotrophic bacteria (Friedmann 1982), with demonstrated heterotrophic respiration at temperatures as low as −20ºC (Goordial et al 2016). Similar colonized lithic substrates are widespread in the Dry Valleys (Cary et al 2010, Cowan et al 2010, De Los Rios et al 2014, Friedmann 1982, Friedmann et al 1988), and typically harbour a relatively high microbial diversity compared to surface soils (Pointing et al 2009), supporting the idea that microbial activity in extremely dry, cold or hot deserts is largely confined to specialized lithic habitats (Pointing and Belnap 2012, Wierzchos et al 2013). The sharp biological contrast between permafrost soils and sandstone rocks suggests that the physical nature of the microenvironment plays a decisive role in the habitability of this extremely cold region, and the potential absence of *in situ* biological activity in the permafrost soils points to a fundamental cold threshold for life, a very rare case on Earth that can be used to constrain the natural cold limit of biological processes.

The objective of this study was to assess the functional differences that underlie the success of the cryptoendolithic communities in comparison to permafrost soils in University Valley, and to assess both for the functional capacity of microorganisms to survive in the extremely hyper-arid, cold and oligotrophic environment. Since carbon fixation is thought to be crucial to cryptoendolithic function, we also assessed the activity and diversity of the photoautotrophic community members in such a hostile environment. The data presented here is the first metagenomic sequencing of Antarctic Dry Valley permafrost completed to date, as well as the only polar cryptoendolith metagenome sequenced to date. We used the permafrost soil

metagenome to compare to other cold or arid soils globally to gain insight into why permafrost in the high elevation Dry Valleys could be inhospitable to life compared to permafrost elsewhere which experience similarly low temperatures, and in which microbial activity has been unambiguously detected (Goordial et al 2013).

Results and Discussion

Metagenome and soil summary

Table 1 shows an overview of the metagenome statistics. The permafrost soils used for metagenomics analysis had a gravimetric moisture content of 9.83%, and was very oligotrophic with 0.02% total carbon, and total nitrogen concentrations below detection limits (<0.001 %). Large amounts of soil were used for DNA extraction (60 g), however due to the low biomass (3 x 10^3 cells g^{-1}), multiple displacement amplification (MDA) was required to generate enough reads. While there are inherent amplification biases known to be associated with MDA (Abulencia et al 2006, Yilmaz et al 2010), the use of MDA was seen as an acceptable compromise to access the functional potential of the very low cell density permafrost soil which was not previously accessible. MDA was not necessary for cryptoendolith samples. Only 63,452 (8.6%) sequences from the permafrost soils and 193,269 (17.4%) sequences from the cryptoendolith metagenomes could be assigned an annotation using the M5 non-redundant protein database (M5nr). With 256,721 annotated reads in this study, we did not identify or capture the entire genetic diversity in these samples.

Microbial community composition

Based on all annotated genes in the metagenomes, the permafrost soil community was Bacteria dominated (86%), with a smaller Eukaryotic fraction (12%), primarily belonging to the fungal phylum *Ascomycota* (Table 2). The cryptoendolith community was mostly Eukaryotic (54%), comprised of the fungal phyla *Ascomycota* (45%) and *Basidiomycota* (2%), as well as the algal phyla *Chlorophyta* and *Streptophyta* (5%), reflecting the lichen dominated community that is prevalent in the Dry Valleys (de la Torre et al 2003, Sun et al 2010). Algae were nearly absent in the permafrost soils (~0.1%), and similarly, photoautotrophic bacteria belonging to C*hloroflexi, Cyanobacteria, and Chlorobi* were detected in small amounts in the cryptoendolith but not in the permafrost soil metagenome.

Similar to other Dry Valleys*, Ascomycota* and *Basidiomycota* were the dominant edaphic and lithobiontic fungal phyla, but the phylum *Chytridiomycota,* abundant in Dry Valley active layer soils (Dreesens et al 2014), was absent in the University Valley permafrost soils and was negligible in the cryptoendolith community (<0.001%). The fungi to bacteria ratio was low in the permafrost soil, as has been observed in the Dry Valleys before, likely due to the low water activity, low C:N ratios, and more extreme conditions that restrict fungal growth and dispersal in high elevation inland soils, while still permitting bacterial survival (Dreesens et al 2014). Though more abundant among the cryptoendolith, similar *Ascomycota* fungi were found in both habitats, and consisted mainly of *Eurotiomycetes*, *Sordariomycetes*, and *Dothideomycetes*. Isolates from these classes are known to be polyextremophillic and are found as parasymbionts (symbionts to lichen) in Antarctic lithic habitats (Selbmann et al 2005). *Eurotiomycetes* and *Dothideomycete*s include the 'black yeast' fungi, which are melanized and are known for their desiccation and UV resistance (Ruibal et al 2009, Selbmann et al 2005), and may play an important role in

community protection from excessive UV radiation, for example by providing an opaque barrier above photobionts in the lithobiontic community (Selbmann et al 2013).

Actinobacteria were the predominant bacterial phylum identified in both the permafrost soil and cryptoendolith metagenomes (64 % and 20% respectively). *Firmicutes, Bacteriodetes and Alpha proteobacteria* were also abundant, similar to other regions in the Dry Valleys (Cary et al 2010, Goordial and Whyte 2014) (Table 2). Desiccation and radiation resistant *Rubrobacteridae* (cryptoendolith: 0.5%, permafrost: 0.01%) and *Deinococcus-Thermus* group bacteria (cryptoendolith: 0.5%, permafrost: 0.02%) commonly found in hot and cold desert soils were however low in the permafrost soil metagenome. These extremophiles are known to be resistant to desiccation, ionizing radiation, UV radiation, and reactive oxygen species (Ferreira et al 1999, Makarova et al 2001, Webb and DiRuggiero 2013). Resistance to the fluctuating moisture conditions, as well to UV radiation would be advantageous in the cryptoendolithic and surface soil habitats, compared with the relatively stable, and dark subsurface permafrost environment. *Deinococcus-Thermus* group bacteria have been previously found to be dominant members of Dry Valley cryptoendolithic communities (de la Torre et al 2003). Negligible Archaeal sequences were identified in both metagenomes (0.4% in the cryptoendolith and 0.01% in permafrost soil) consistent with previous reports that Archaea are absent, or difficult to detect in Dry Valley soils and lithobiontic communities (Lee et al 2012, Pointing et al 2009). The most abundant archaeal classes were *Methanomicrobia* and *Halobacteria,* which were found to represent 0.1% and 0.09% of reads respectively in the cryptoendolithic community.

Viral reads were a minor component of the cryptoendolith metagenome (0.5%), but comprised a relatively large proportion of the permafrost soil metagenome (2%). Viral families identified were consistent with those detected in other Dry Valley soils and lithic environments (Wei et al , Zablocki et al 2014), and *Microviridae* and *Siphoviridae* were the most abundant in both metagenomes (Table S *1*). *Microviridae* and *Siphoviridae* are known to infect bacteria, and have been found associated with *Arthrobacter, Streptomyces, Staphylococcus* and *Bacillus* species in Antarctic soil (Hopkins et al 2014, Swanson et al 2012), genera also identified in University Valley soils (Goordial et al 2016). The comparative paucity of viral reads in the cryptoendolithic community is in contrast to previous comparisons of lithic habitats and open soils in the Dry Valleys, which found that lithic habitats harbour a higher abundance and diversity of viruses (Zablocki et al 2014). The reasons underlying these differences are unknown, as little is yet known about viral roles in community ecology in the Dry Valleys. In Arctic active layer soils, viruses have been demonstrated to exert a top down control on soil communities, decreasing both biomass and activity (Allen et al 2010). The proportion of viral reads found in this study are higher than those observed in the metagenomes of permafrost soils in the Arctic and Dry Valley surface soils (Table S3) which range from 0.008% to 0.09% of total reads.

Community fingerprinting data for individual cryptoendolith and ice-cemented permafrost soil samples (Figure S1) showed that samples that community composition at the phylum level had little variability in University Valley soils, or the cryptoendoliths samples here. Cryptoendoliths were dominated by reads assigned as cyanobacteria or chloroplasts- metagenomic sequencing, in conjunction with the plastid sequencing efforts described below indicate that most of these sequencings are from algae, with small cyanobacterial populations. Permafrost samples were

dominated by gamma-proteobacteria primarily belonging to the orders Alteromonadales, Oceanospirallales, and Pseudomonadales.

Dry Valley surface soil communities have been found to be highly localized, suggesting a high degree of endemism within each valley (Lee, 2012), and indicating that aeolian input of microorganisms by strong katabatic winds throughout the Dry Valleys may play a more limited role in community composition compared to local conditions. For example, a recent 16S rRNA gene survey of aerosols in the lower elevation Dry Valleys showed few OTUs in common with the nearby surface soils (Bottos et al 2014). In University Valley, soils are largely derived from the weathering and erosion of the colonized valley walls (Heldmann et al 2013, Tamppari et al 2012), and a previous molecular survey found that the cryptoendoliths and surface soils share few OTU's in common (Goordial et al 2016). It is likely that the permafrost soil community in University Valley is derived from a mixture of wind deposited cells and weathered cryptoendoliths, in which subsequently only the few cells that can form spores, remain dormant, or have advantageous adaptations to the extremely oligotrophic, arid and cold permafrost environment, may survive.

Functional diversity in University Valley cryptoendolith and permafrost communities

The most abundant genes in both metagenomes were related to housekeeping functions such as carbohydrate metabolism, amino acids and derivatives, protein metabolism, respiration, and co-factor, vitamin and pigment production (Figure 3). Genes responsible for the degradation of cellular material were more abundant in the permafrost soil and included several involved with

murein recycling, and N-acetylglucosamine and chitin utilization. These genes could be advantageous for using cellular material as a nutrient source, including potential biomass from eroding cryptoendolithic communities. The cryptoendolith metagenome was enriched with genes involved with photosystems, CO_2 fixation, and auxin biosynthesis, phytohormones which can stimulate growth and production of antioxidants in algae (Piotrowska-Niczyporuk and Bajguz 2014). Genes reflective of the dense microbial consortia living within the narrow colonized zone were found in the cryptoendolith metagenome including genes associated with quorum sensing (N-acyl homoserine lactone hydrolase, S-adenosylmethionine synthetase), multidrug efflux pumps, antibiotic resistance (penicillin, fluoroquinolones, methicillin, vancomycin) and genes for the production of secondary metabolites known to be antibacterials and antifungals (phenanzine, clavulanic acid). Biosynthesis genes for a number of cofactors, vitamins and prosthetic groups which can support photosynthesis were present in both metagenomes (coenzyme B12, thiamine, biotin). Both metagenomes had the functional potential for catabolism of a diversity of aromatic compounds, poly- and oligosaccharides and carbohydrates (e.g. catabolism of benzoate, catechol, gentisate, maltose, mannose, xyloglucan, lactose).

Stress Response and cold adaptation

The cryptoendolith metagenome had both a higher relative abundance (Figure 3) and higher diversity of stress response genes (measured as number of different stress response genes) compared to the permafrost soils, with 87 and 34 stress response genes respectively (Figure S2). The two metagenomes only shared 25 of the 96 stress response related genes detected. Known cold-adaptive genes were found in both the soils and cryptoendolith metagenome. General

247 microbial adaptations to cold environments include mechanisms that protect the cell from
248 freezing, preserve enzymatic membrane function, protect against reactive oxygen species (ROS),
249 and protect against osmotic stress caused by the increasing salt and solute concentrations as
250 water freezes (Goordial et al 2013). The stress response pathways in both habitats represented
251 redundant functions, mostly associated with the osmotic and oxidative stresses, which are
252 characteristic of cryoenvironments. For example, shared proteins in both habitats included those
253 involved with glycerol uptake, and proline and glycine betaine transport across membranes, these
254 are cryoprotectants and compatible solutes which are commonly used by psychrophilic
255 microorganisms as a strategy to cope with osmotic stress in sub-zero environments (Methé et al
256 2005, Mykytczuk et al 2013). Cold-shock proteins were found in both metagenomes, although
257 bacterial antifreeze protein, which prevents ice-crystal formation, was only found in the
258 cryptoendolith. Other shared stress response genes were related to general stress response
259 functions like chaperones, sigma B stress response, carbon starvation or phage shock protein A
260 (pspA); phage shock protein A is a stress response gene involved in maintaining cell membrane
261 and proton motive force integrity and is induced during extremes of temperature, osmotic stress,
262 and filamentous phage infection.
263
264 *Nutrient Cycling*
265
266 In highly oligotrophic soils like those encountered in University Valley, carbon and nutrient
267 sequestration is important. Examining the presence and absence of metabolic pathways we found
268 no evidence for functioning autotrophic pathways in the permafrost soil as determined by the
269 absence of key enzymes in the Calvin Benson cycle, the reductive TCA cycle, the reductive

acetyl-coA pathway, and the hydroxypropionate cycle carbon fixation pathways (Table 3). The paucity of genes associated with autotrophy suggests these soils are dependent on heterotrophic substrates. As expected, the cryptoendolith metagenome contained the genes for CO_2 fixation with the Calvin Benson cycle. Genes associated with metabolism of trace gases and other C1 compounds (methanogenesis, acetogenesis, methanotrophy) were limited or absent in both metagenomes, with the exception of methane monooxygenase (*mmoX*) detected in the cryptoendolith metagenome. Genes required for heterotrophy were abundant in both metagenomes including genes required for acetate metabolism, a compound not mineralized at sub-zero temperatures in University Valley soils (Figure 2 and (Goordial et al 2016)). Key genes in the glyoxylate pathway were found in both metagenomes, including isocitrate lyase and malate synthase, and it is possible the CO_2 releasing steps of the TCA cycle can be bypassed, although heterotrophic activity as inferred from respired $^{14}CO_2$ was detected in University Valley permafrost soil microcosms at 5ºC (Figure 2) so it is unlikely that the glyoxylate pathway is responsible for the lack of microbial activity observed at sub-zero temperatures. Previous analysis of University Valley permafrost soils indicated that due to the low salt concentration, the amount of liquid water at below freezing temperatures is limited only to thin films adhering to sand grains (Goordial et al 2016). The activity observed only above freezing temperatures may reflect water newly available for cells, which would be otherwise dormant at *in situ* freezing temperatures.

Nitrogenase reductase (*nifH*) sequences were not detected in either metagenome. Cryptoendolithic communities in the Upper Dry Valleys largely lack the ability to fix nitrogen (as measured by acetylene reduction), likely because of available nitrates which are

atmospherically deposited and which have low leaching rates in desert environments (Friedmann and Kibler 1980). Some nitrogen cycling genes (nitrate and nitrite reductases) were recovered from both metagenomes, although other denitrification genes (nitric oxide reductase and nitrous oxide reductase) were absent. The lack of nitrogen and carbon fixation capacity in the permafrost soils separates these soils from Arctic permafrost and lower elevation (<1000 m.a.s.l.) Dry Valleys surface soils where both photoautotrophic and diazotrophic pathways have been identified in functional microarray and PCR surveys (Chan et al 2013, Niederberger et al 2012, Yergeau et al 2010).

Microbial activity in University Valley and characterization of the photoautotrophic cryptoendolith community

Photoautotrophic microorganisms drive carbon acquisition in the sandstone cryptoendoliths, and may be a source of organic matter to the permafrost soils, and thus could play a key role in ecosystem function. In order to get better resolution of the photoautotrophic diversity we carried out 454 pyrosequencing on two cryptoendolith samples targeting the 23S rRNA plastid gene found in photosynthetic organisms, including cyanobacteria and algae (Sherwood and Presting 2007). The cryptoendolith photoautotroph community was dominated almost entirely by *Trebouxia*, an algae known to form lichenizing associations, which made up over 99% of sequences. Lichen dominated cryptoendoliths in the high elevation Dry Valleys have been previously shown to be mono-specific (de la Torre et al 2003), but we found a high diversity of *Trebouxia sp*. in the cryptoendolith samples, with 365 OTU's (97% cut-off) for this genus between both cryptoendolith samples, which only shared 24 OTU's in common (Figure S3).

316 *Cyanobacteria* were a minor component made up of 4 OTUS's (representing 33 sequences) in
317 one cryptoendolith sample, and were absent in the other (Figure S3). Other photosynthetic
318 organisms were not detected using pyrosequencing, including *Chloroflexi*, and the *Streptophyta*
319 algae annotated in the metagenome.
320
321 We were able to isolate green algae from the cryptoendolith samples (Table S ***2***) belonging to the
322 genus *Stichococcus* and *Desmococcus*. Photoautotrophs could not be cultured from permafrost
323 soils using the same methodologies. Isolates identified as *Stichococcus* EN2JG and *Desmococcus*
324 EN5JG were adapted for cold temperatures and demonstrated growth (Table S***2***) and chloroplast
325 autofluorescence at −5ºC (Figure 2). Notably, the observed growth occurred with no media
326 amendments to prevent cultures from freezing, indicating these isolates are synthesizing freezing
327 point depressants to maintain a liquid culture at sub-freezing temperatures. The isolates were not
328 capable of growth when glycerol (5%) or NaCl (5%) were added as freezing point depressants,
329 and the liquid media tested here froze at the other temperatures tested (< −10ºC), thus potential
330 growth at lower temperatures could not be measured. Two *Stichococcus* isolates differed in their
331 growth characteristics; *Stichococcus* sp. EN2JG was a eurypsychrophile with an optimal
332 temperature of 22ºC and a minimum temperature of −5ºC, while *Stichococcus* sp. UV2BC was a
333 stenopsychrophile incapable of growth at 22ºC, with an optimal temperature of 10ºC and a
334 minimum temperature of 0 ºC. Differing growth optima may occur in the diverse unculturable
335 algae surveyed here as well, and would result in communities which could fix carbon over the
336 breadth of fluctuating conditions the sandstone cryptoendoliths would experience.
337

We also carried out Pulse Amplitude Modulated (PAM) fluorometry to determine the activity of the photosystem II [PS(II)] of phototrophic members of the cryptoendoliths. Significant PS(II) activity (measured as variable fluorescence, (Fo - Fm)/Fm) was measured to be 0.618, 0.560, 0.467 at 20ºC, 0ºC and −20ºC respectively, indicating that colder temperatures affected PS(II) efficiency, but photosynthesis could still potentially occur at −20ºC. We found that similar amounts of amounts ^{14}C labelled acetate was mineralized at 5ºC (4.1%), −5ºC (3.8%), and −10ºC (4.4%) over 100 days (Figure 2); thus both the heterotrophic and photoautotrophic communities within the cryptoendoliths display thermal plasticity allowing activity over a range of temperatures that overlaps with those observed in the natural environment during the summer months. In contrast, heterotrophic activity in permafrost soils was undetectable at −5ºC and -10ºC, and could only be detected at 5ºC, a temperature which is not encountered *in situ* and likely reflects the activation of dormant but viable cells.

Comparison of University Valley permafrost with other desert and permafrost metagenomes

An ordination (Figure 4) was created to examine the functional similarities and differences of University Valley permafrost soil with other permafrost and desert environments. A list of the metagenomes used for comparison in this study is available in Table S3. The University Valley permafrost soil metagenome clustered most closely with other permafrost metagenomes from the Arctic, and separately from the more geographically proximate Dry Valley active layer soils. This may indicate that the permafrost soils in University Valley are more similar to Arctic permafrost than previously assumed based on the low biomass and lack of microbial activity

previously detected (6). Compared to the hot and cold desert soils, the permafrost metagenomes were enriched in genes associated with osmotic stress, which would be advantageous in the brine veins thought to exist within permafrost as a potential microbial habitat, where salts, solutes and microorganisms could be concentrated together during freezing in a similar manner to sea ice (Junge et al 2001). The permafrost soils metagenomes were also enriched in integrases and transposases, and antibiotic and antiseptic resistance genes including beta-lactamases, vancomycin and acriflavin resistance. It is not known what role these genes would have, though it is possible that in permafrost soils microbial biomass becomes concentrated within brine veins, increasing microbial competition for limited nutrients and possibly occurrences of lateral gene transfer. University Valley permafrost soil was an outlier to the permafrost samples in this respect and had the lowest proportion of antibiotic resistance genes, integrases and transposases (Figure S4). The contrast could be due to a combination of unique factors in University Valley soils; the low soil salinity is prohibitive to the formation of brine veins where cells could concentrate (Goordial et al 2016), biomass in University Valley permafrost soils is extremely low (10^3 cells g^{-1}), and as indicated by the absence of metabolic activity, microorganisms in the permafrost soils are likely not competing, but are dormant. University Valley permafrost soils were less functionally equipped with oxidative stress, general stress response and cold shock genes, though intriguingly had the highest proportion of phage related genes (mostly phage capsid proteins) compared to the other permafrost metagenomes. Rather than a diversity of stress response functions, University Valley had a comparatively high proportion of a number of genes associated with sporulation and spore DNA protection (Figure S4). While survival on long time scales is important in all permafrost environments (Figure 4), traits which allow cells to persist in permafrost, rather than for growth or activity, are especially emphasized in University Valley

permafrost soils where the conditions may be too extreme for the activity of even cold adapted extremophiles.

Comparison of University Valley cryptoendolith metagenome with other photoautotroph based metagenomes

To our knowledge, this is the first terrestrial cryptoendolithic metagenome reported to date, and no metagenomes for related habitats such as hypoliths and chasmoendoliths are currently available in public databases.. Thus we are limited in our ability to compare the University Valley cryptoendolith to lithic environments in other hot or cold deserts. We chose to compare the biofilm like cryptoendolith community with other communities which have a large photoautotroph component, and included in our ordination metagenomes from an Alpine lichen community, polar microbial mats, and a glacial cryoconite hole microbial community. The cryptoendolith metagenome did not cluster strongly with any of these metagenomes. The cryptoendolith shared with the lichen metagenome a higher proportion of genes involved with quorum sensing, and cofactor, vitamin and pigment production, a reflection of the symbiotic relationship between mycobionts and phycobionts seen in both the cryptoendolith and lichen communities. The genes shared with the Antarctic microbial mat metagenome were important in biofilms, including those involved in adhesion, extracellular polysaccharides, and siderophore production; in the cryptoendolith these traits would be useful in rock colonization, and iron acquisition/mobilization functions which result in the characteristic red banding pattern seen in cryptoendolithic communities (Figure 1). Phages, bacterial cytostatic and antibiotic production and resistance was most abundant in the cryptoendolith metagenome, possibly indicative of a

higher level of microbial competition and predation than in the other metagenomes used here for comparison.

Conclusion: Habitability conditions in University Valley permafrost soils and lithic habitats.

It has already been postulated that while microorganisms are present in the permafrost soils (Goordial et al 2016, Tamppari et al 2012), microbial activity is likely non-existent in parts of University Valley where soils are permanently cryotic (Goordial et al 2016). We refer to these soils as non-habitable, but not sterile, and the resulting permafrost soil community is likely a mixture of aeolian and cryptoendolithic origin. This interpretation is supported by the metagenomic data presented in this study, which revealed less cold and general stress response functional diversity, critical for life in permafrost soils, whereas sporulation (i.e. dormancy) is an emphasized function. The functional potential for recycling of cellular material, as well as the large presence of phage associated genes suggest that if there is an active component of University Valley permafrost soils, it could survive using scavenged organic matter, possibly of endolithic origin since that is the only relevant source of biomass in the valley.

On the other hand the cryptoendolithic communities that colonize the valley walls appear to be adapted to the harsh conditions within the valley (Figure 5), as evidenced by the development of a complete ecological community, including photoautotrophic algae and bacterial/fungal consumers that are viable and active over the range of temperatures the cryptoendoliths experience, and by the diversity of stress response functions and nutrient cycling pathways. We

have added to the functional knowledge of lithic communities which are known 'hot spots' of productivity in cold and dry environments, including evidence for the likely presence of community competition in addition to the well-known symbiotic interactions, as indicated by the presence of antifungal and antibacterial production and resistance genes.

The stark biological contrast between permafrost soils and lithobiontic habitats is due largely to the physical properties of the lithic substrate. Primarily, the sandstone favours the occurrence of wet events through inducing the melting of snow (Friedmann 1978, Friedmann et al 1987). Once wet, surface tensions between thin films of water and the rock matrix slow down evaporation, and extends the window for metabolic activity (Friedmann et al 1987). This, together with the protection from UV radiation while still allowing for photosynthetic activity, represent decisive survival advantages that ultimately control habitability under extreme cold and dry conditions.

Our results evidence that caution should be taken when interpreting function solely from genomic analyses, which cannot differentiate between vegetative, dormant and dead cells, especially in stable and cold permafrost soils which are likely highly preserving for nucleic acids. Additionally, the lack of detection of genes found in other Dry Valley environments but not in University Valley permafrost soils (antifreeze proteins, nitrogenase genes etc) may be due to limitations in depth and coverage, the small sample size in this study, as well as the biases introduced by MDA; future metagenomic studies in the Dry Valleys will likely overcome these drawbacks as sequencing technologies improve and lower in cost. Metagenomic analysis is best complimented by functional validation and activity assays, though given the difficulties in culturing and isolating organisms from such extreme environments, metagenomics sequencing is

453 a good proxy for the functional potential of environments which may otherwise be inaccessible.
454 Future studies utilizing transcriptomic, proteomic and activity assays targeting some of the
455 functions identified in this study are the next step to understanding how microbial communities
456 are adapted to thrive and survive in one of the coldest and driest terrestrial habitats on Earth.
457

458 **Acknowledgments**

459 This work was supported by NASA's ASTEP program and with field support via NSF/OPP
460 (project B-302-M). Support was provided by the Natural Sciences and Engineering Research
461 Council (NSERC) Discovery Grant Program, NSERC Northern Supplements Program, and
462 NSERC CREATE Canadian Astrobiology Training Program (CATP). Metagenomic data sets
463 are publically available at MG-RAST (4594281.3, 4594282). Pyrosequencing data has been
464 deposited in the NCBI Sequence Read Archive (SRA) under project PRJNA290089.
465

Material and Methods

Sample collection and preparation

The University Valley permafrost core sample used in this study was collected in the 2009 summer field season, located at 77d 51.870s S, 160d43.524s E (elevation 1700 m.a.s.l). Depth from the surface to the ice-cemented ground was 22 cm. An 18 cm ice-cemented permafrost core was collected with a SIPRE corer along with overlaying dry permafrost Samples were shipped to McGill University in a thermally insulated box and maintained at −20ºC until processing. Initial core processing took place in a walk-in freezer held at −5ºC, in a laminar flow hood where 1 cm of the outside of the core was removed with a sterilized chisel. An additional 1 cm of the outside core was removed in a laminar flow hood at room temperature immediately prior to samples being weighed and aliquoted for analysis. Cryptoendolith samples used in this study were collected in the 2013 summer field season, from Beacon supergroup sandstone boulders located on the South-East facing valley walls. Samples were aseptically collected and maintained at −20ºC until processing.

Soil Analysis

The soils were analyzed for total carbon and total nitrogen by combustion at 900°C with a Carlo Erba Flash EA 1112 NC Soils Analyzer which has an analytical error of ±1%. Gravimetric moisture content was measured as a percentage of dry weight. 20 g of soil was oven dried at 100°C for 48 hours and weighed using a Mettler AE 163 analytical balance with an accuracy ±0.02 mg. The pH of soils was measured using a 1:2 slurry of soil:deionized water with a Fisher Scientific pH electrode (Fisher Scientific), with an efficiency slope of >95%.

DNA extraction

The top ten 10 cm of the ice-cemented permafrost core (22-32 cm depth from the surface) was used for metagenomics analysis. Community DNA was extracted from 2 g of permafrost soil using the UltraClean Soil DNA Isolation kit (MoBio Laboratories Inc., Carlsbad, California, USA), as described in the alternative protocol for maximum yield, and a bead beating step was added to aid lysis. 30 extractions (60 g total permafrost soil) were performed and the resulting DNA was pooled and concentrated. DNA from cryptoendoliths was extracted from 6 grams total of crushed rock using the same DNA extraction protocol as for permafrost. Negative controls (H_2O in place of sample) underwent identical handling during the extraction procedure and were used as templates for PCR using 16S rRNA gene primers (27F and 1492R) to ensure no contamination during extraction.

Metagenomic Sequencing and Analysis

DNA was sent to Molecular Research LP (Shallowater, Texas, USA) for sequencing. There, the library was prepared using Nextera DNA Sample preparation kit (Illumina) following the manufacturer's user guide. Both the samples were first purified using PowerClean DNA Clean-up Kit (MoBio) and concentration of purified gDNA was evaluated using the Qubit dsDNA HS Assay Kit (Life Technologies). Because of low DNA concentration for the ice-cemented permafrost sample Multiple Displacement Amplification (MDA) was performed at 30°C for 16h using the REPLI-g Midi Kit (Qiagen) according to the manufacturer's instructions for 2.5uL of input DNA. Once amplification was complete, the concentration of the sample was again determined and each sample was diluted accordingly to achieve the recommended DNA input of 50ng at a concentration of 2.5ng/µL. Subsequently, the sample underwent the simultaneous fragmentation and addition of adapter sequences. These adapters are utilized during a limited-cycle (5 cycles) PCR in which unique index was added to the sample. Following the

library preparation, the final concentration of the library was measured using the Qubit dsDNA HS Assay Kit (Life Technologies), and the library size was determined using the Experion Automated Electrophoresis Station (Bio-Rad). The libraries (12.5 pM) were pooled together and was sequenced by using 600 Cycles v3 Reagent Kit (Illumina) in MiSeq (Illumina). Sequences were processed using MG-RAST (Meyer et al 2008), artificial replicate sequences produced by sequencing artifacts were removed (Gomez-Alvarez et al 2009), and sequences were quality trimmed using the default settings for dynamic trimming (sequences contain <5 bp below a phred score of 15). Any human or chordata contaminants were removed from the dataset. To determine the presence or absence of functional genes, we used GenBank annotated proteins in MG-RAST (e-value $\leq 10^{-5}$, alignment length >15). We used the statistical probability model in Statistical Analysis of Metagenomic Profiles (STAMP) (Parks and Beiko 2010) (version 2.08; Faculty of Computer Science, Dalhousie University) to identify the biologically relevant differences between the permafrost and cryptoendolith metagenomes. A pairwise statistical comparison of the two metagenomes analyses was carried out using clustering based SEED subsystem annotations ($E \leq 10^{-5}$, similarity > 60 %, alignment length >15), using a two-sided Chi-square test (with Yates) statistic with the DP: asymptotic-CC confidence interval method and the Bonferroni multiple test correction. A *P*-value of <0.05 was considered significant, and an effect size filter for ratio of proportions (RP) effect size <2.00 and a difference of proportions of <2.5. Comparisons of the University Valley metagenomes were made to other publically available metagenomes from similar hot and cold deserts, or microbial mat communities (Table S 3), all metagenomics data was processed through MG-RAST to make analyses comparable. Relative abundance was used to calculate Bray-Curtis distances between sample pairs using the "vegdist" function of the "vegan" package (http://vegan.r-forge.r-project.org/) in Rstudio (version

534 0.98.1091). Principle coordinate analyses (PCoA) analyses were performed using the 'cmdscale'
535 function. Relative abundance of level 1 of the SEED hierarchy were superimposed on the
536 ordination using the "envfit" function.

537 **Pyrosequencing of plastid gene amplification**

538 DNA from the cryptoendoliths was sent for pyrosequencing analyses at the Research and
539 Testing Laboratory (Lubbock, TX, USA) using the Roche 454 GS-FLX platform (Roche 454,
540 Branford, CT, USA). Sample libraries of partial bacterial/ algal 23S rRNA amplicons were
541 produced using the forward primer (5'GGACAGAAAGACCCTATGAA-3') and reverse primer
542 (5'-TCAGCCTGTTATCCCTAGAG- 3') that flank the V domain of the 23S plastid rRNA gene.
543 Data was processed using Mothur (Schloss et al., 2009), Briefly, sequences were quality filtered
544 by removing primer sequences, reads < 150 bp long, sequences with ambiguous base calls, and
545 homopolymer repeats greater than 8bp. Chimera removal using chimera.uchime within Mothur
546 was used to further reduce sequencing error prior to alignment and clustering. A total of 13,557
547 sequences were analyzed after quality control. Sequences were aligned to the Silva LSU bacterial
548 database (Accessed March 2015) and OTUs were clustered using average-neighbour clustering
549 with a 97% cut-off. The 'get.oturep' command of Mothur was used to retrieve a representative
550 sequence for each OTU. Representative sequences were classified using the MEGAN5 software
551 (v. 5.3.0) (Huson et al 2007) after BLASTn searches against the GenBank nt database
552 (http://www.ncbi.nlm.nih.gov/GenBank/) (accessed June 2014) with default settings, and by
553 excluding noncultured/environmental sequences from the target database. For MEGAN5
554 classification, LCA parameters were changed from default as to favour the taxonomic
555 information of the best BLASTn hits to be assigned to a given read; LCA parameters were set to
556 'Min Support: 2', 'Min Score: 100', 'Top percent: 2', and 'Min complexity:0'. The primers used

in this study were found to amplify some non-phototrophic bacteria (belonging to acidiphilum, and caulobacter) representing <0.01% of reads and which were manually removed from the dataset.

Community fingerprinting of cryptoendoliths and soil samples

Environmental DNA was extracted from the colonization zone of 4 individual sandstone rocks or from 0.2 g of ice-cemeted permafrost soil from 5 individual samples using the PowerSoil DNA isolation kit (MoBio laboratories Inc., Solana Beach, CA). Permafrost samples were from 3 separate cores and at varying depth: core 2, 1 cm; core 4, 1 cm, 3 cm, 5 cm; and core 14, 2 cm. DNA was amplified using the barcoded universal primers 338F and 806R for the V3–V4 hypervariable region of the 16S rRNA gene and amplicons from 3 reactions were pooled together for sequencing using the Illumina MiSeq platform. The QIIME package (v1.6.0) was used for quality control with following criteria: 1) minimum and maximum length of 200 bp and 400 bp; 2) an average of q25 over a sliding window of 25 bp. If the read quality dropped below q25 it was trimmed at the first base pair of the window and then reassessed for length criteria; 3) a perfect match to a barcode sequence; 4) a match to E. coli 16S rRNA gene and 5) presence of the 16S primer sequence used for amplification. Sequences were binned based on sample-specific barcode sequences and trimmed by removal of the barcode and primer sequences (forward if present and reverse). Chimera removal using chimera.uchime within Mothur was used to further reduce sequencing error prior to alignment and clustering. Sequences were aligned to the Silva reference files provided by Morthur (release 119) (Accessed March 2016) and OTUs were clustered using average-neighbour clustering with a 97% cut-off.

580

581 **Heterotrophic Radiorespiration Assay**

582 5 g of permafrost was added to individual microcosms as Steven et al. 2007. Each
583 microcosm was performed in triplicate, and included triplicate sterilized controls (autoclaved
584 twice for 2 hours at 120°C and 1.0 atm, with a 24 h period between autoclavings). Microcosms
585 were spiked with 0.045 mCi ml^{-1} (~100,000 disintegrations per minute) of 1-^{14}C acetic acid.
586 Cold acetic acid was added to a final concentration of 15 mM acetic acid per microcosm in a
587 total volume of 40 µl. The CO_2 trap consisted of 1 M KOH for microcosms incubated at 5°C,
588 −5°C and 1 M KOH + 20% v/v ethylene glycol for microcosms incubated −10°C and −15°C. For
589 cryptoendolithic microcosms 3 g of crushed rock from the visibly colonized area of the
590 sandstone was used for each microcosm, and spiked with cold and radioactive acetate as
591 described. Measurements of radioactivity were determined by liquid scintillation spectrometry
592 on a Beckman Coulter (CA, USA) LS 6500 Multi-purpose Scintillation Counter.

593 **Pulse Amplitude Modification PAM methodology**

594 Chlorophyll *a* fluorescence was measured with a PAM fluorometer (WATER-PAM,
595 Heinz Walz GmbH). After 30 minutes dark adaptation, initial fluorescence ($F_{O)}$ was measured
596 and represents the point where all PSII reaction centres are open and the most light energy can be
597 used for photochemistry rather than being emitted as fluorescence. The sample was then given a
598 saturation pulse until all reaction centres were closed and electron acceptors saturated, all light
599 energy is given off as maximal fluorescence in this state (F_M). Photosynthetic efficiency of
600 photosystem II (PSII) was measured as F_V/F_M, where variable fluorescence (F_V) is calculated as
601 the difference between initial fluorescence (F_O) and maximal fluorescence (F_M).

602 **Isolation and characterization of photoautotrophs**

1g of permafrost soil, or of the colonized band of the cryptoendolith was sampled, homogenized and added to a sterile tube containing 3mL of 0.1% sterile sodium pyrophosphate and 0.5g of glass beads. Following 1 min of vortexing, 100µL of the suspended cell solution was used to inoculate liquid media and agar plates of BG11, CHU-10 and SNAX media. Plates and liquid enrichment cultures were incubated at 5°C and 20°C, in the presence of 24 hours 6400K full spectrum light (T5HO bulb, Sunblaster) until growth was observed. Isolates were then characterized for growth at −5°C, 0°C, 5°C and 20°C, as well as with 5% NaCl and 5% glycerol added to media.

References

Abulencia CB, Wyborski DL, Garcia JA, Podar M, Chen W, Chang SH *et al* (2006). Environmental whole-genome amplification to access microbial populations in contaminated sediments. *Appl Environ Microb* **72:** 3291-3301.

Aislabie JM, Chhour K-L, Saul DJ, Miyauchi S, Ayton J, Paetzold RF *et al* (2006). Dominant bacteria in soils of Marble Point and Wright Valley, Victoria Land, Antarctica. *Soil Biology and Biochemistry* **38:** 3041-3056.

Allen B, Willner D, Oechel WC, Lipson D (2010). Top-down control of microbial activity and biomass in an Arctic soil ecosystem. *Environmental Microbiology* **12:** 642-648.

Bakermans C, Skidmore ML, Douglas S, McKay CP (2014). Molecular characterization of bacteria from permafrost of the Taylor Valley, Antarctica. *FEMS Microbiology Ecology* **89:** 331-346.

Bottos EM, Woo AC, Zawar-Reza P, Pointing SB, Cary SC (2014). Airborne Bacterial Populations Above Desert Soils of the McMurdo Dry Valleys, Antarctica. *Microbial Ecology* **67:** 120-128.

Campbell IB, Claridge G (1987). *Antarctica: Soils, Weathering Processes and Environment: Soils, Weathering Processes and Environment*, vol. 16. Elsevier.

Cary SC, McDonald IR, Barrett JE, Cowan DA (2010). On the rocks: the microbiology of Antarctic Dry Valley soils. *Nature Reviews Microbiology* **8:** 129-138.

Chan Y, Van Nostrand JD, Zhou J, Pointing SB, Farrell RL (2013). Functional ecology of an Antarctic dry valley. *Proceedings of the National Academy of Sciences* **110:** 8990-8995.

Cowan DA, Russell NJ, Mamais A, Sheppard DM (2002). Antarctic Dry Valley mineral soils contain unexpectedly high levels of microbial biomass. *Extremophiles* **6:** 431-436.

Cowan DA, Khan N, Pointing SB, Cary SC (2010). Diverse hypolithic refuge communities in the McMurdo Dry Valleys. *Antarctic Science* **22:** 714-720.

de la Torre JR, Goebel BM, Friedmann EI, Pace NR (2003). Microbial Diversity of Cryptoendolithic Communities from the McMurdo Dry Valleys, Antarctica. *Appl Environ Microb* **69:** 3858-3867.

De Los Rios A, Wierzchos J, Ascaso C (2014). The lithic microbial ecosystems of Antarctica's McMurdo Dry Valleys. *Antarctic Science* **26:** 459-477.

De Maayer P, Anderson D, Cary C, Cowan DA (2014). Some like it cold: understanding the survival strategies of psychrophiles. *EMBO reports***:** e201338170.

Dreesens LL, Lee CK, Cary SC (2014). The Distribution and Identity of Edaphic Fungi in the McMurdo Dry Valleys. *Biology* **3:** 466-483.

Ferreira AC, Nobre MF, Moore E, Rainey FA, Battista JR, da Costa MS (1999). Characterization and radiation resistance of new isolates of Rubrobacter radiotolerans and Rubrobacter xylanophilus. *Extremophiles* **3:** 235-238.

Friedmann EI (1978). Melting snow in the dry valleys is a source of water for endolithic microorganisms. *Antarctic Journal of the United States* **13:** 162-163.

Friedmann EI, Kibler AP (1980). Nitrogen economy of endolithic microbial communities in hot and cold deserts. *Microbial ecology* **6:** 95-108.

Friedmann EI (1982). Endolithic Microorganisms in the Antarctic Cold Desert. *Science* **215:** 1045-1053.

Friedmann EI, McKay C, Nienow J (1987). The cryptoendolithic microbial environment in the Ross Desert of Antarctica: Satellite-transmitted continuous nanoclimate data, 1984 to 1986. *Polar Biology* **7:** 273-287.

Friedmann EI, Hua M, Ocampo-Friedmann R (1988). 3.6 Cryptoendolithic Lichen and Cyanobacterial Communities of the Ross Desert, Antarctica. *Polarforschung* **58:** 251-259.

Gilichinsky DA, Wilson GS, Friedmann EI, Mckay CP, Sletten RS, Rivkina EM *et al* (2007). Microbial Populations in Antarctic permafrost: Biodiversity, state, age, and implication for astrobiology. *Astrobiology* **7:** 275-311.

Gomez-Alvarez V, Teal TK, Schmidt TM (2009). Systematic artifacts in metagenomes from complex microbial communities. *ISME J* **3:** 1314-1317.

Goordial J, Lamarche-Gagnon G, Lay C-Y, Whyte L (2013). Left Out in the Cold: Life in Cryoenvironments. *Polyextremophiles*. Springer. pp 335-363.

Goordial J, Whyte L (2014). Microbial Life in Antarctic Permafrost Environments. *Antarctic Terrestrial Microbiology*. Springer Berlin Heidelberg. pp 217-232.

Goordial J, Davila A, Lacelle D, Pollard W, Marinova MM, Greer CW *et al* (2016). Nearing the cold-arid limits of microbial life in permafrost of an upper dry valley, Antarctica. *ISME J*.

Heldmann JL, Pollard W, McKay CP, Marinova MM, Davila A, Williams KE *et al* (2013). The high elevation Dry Valleys in Antarctica as analog sites for subsurface ice on Mars. *Planetary and Space Science* **85:** 53-58.

Hopkins DW, Swanson MM, Taliansky ME (2014). What Do We Know About Viruses in Terrestrial Antarctica? In: Cowan DA (ed). *Antarctic Terrestrial Microbiology*. Springer Berlin Heidelberg. pp 79-90.

Hultman J, Waldrop MP, Mackelprang R, David MM, McFarland J, Blazewicz SJ *et al* (2015). Multi-omics of permafrost, active layer and thermokarst bog soil microbiomes. *Nature* **521:** 208-212.

Huson DH, Auch AF, Qi J, Schuster SC (2007). MEGAN analysis of metagenomic data. *Genome Research* **17:** 377-386.

Junge K, Krembs C, Deming J, Stierle A, Eicken H (2001). A microscopic approach to investigate bacteria under in situ conditions in sea-ice samples. *Ann Glacio* **33:** 304-310.

Lacelle D, Lapalme C, Davila A, Pollard W, Marinova MM, Heldmann J *et al* (2015). Solar Radiation and Air and Ground Temperature Relations in the Cold and Hyper-Arid Quartermain Mountains, McMurdo Dry Valleys of Antarctica. *Permafrost and Periglacial Processes*.

Lee CK, Barbier BA, Bottos EM, McDonald IR, Cary SC (2012). The Inter-Valley Soil Comparative Survey: the ecology of Dry Valley edaphic microbial communities. *ISME J* **6:** 1046-1057.

Legendre M, Bartoli J, Shmakova L, Jeudy S, Labadie K, Adrait A *et al* (2014). Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology. *Proceedings of the National Academy of Sciences* **111:** 4274-4279.

Mackelprang R, Waldrop MP, DeAngelis KM, David MM, Chavarria KL, Blazewicz SJ *et al* (2011). Metagenomic analysis of a permafrost microbial community reveals a rapid response to thaw. *Nature* **480:** 368-371.

Makarova KS, Aravind L, Wolf YI, Tatusov RL, Minton KW, Koonin EV *et al* (2001). Genome of the Extremely Radiation-Resistant Bacterium Deinococcus radiodurans Viewed from the Perspective of Comparative Genomics. *Microbiology and Molecular Biology Reviews* **65:** 44-79.

Marchant DR, Head JW (2007). Antarctic dry valleys: Microclimate zonation, variable geomorphic processes, and implications for assessing climate change on Mars. *Icarus* **192:** 187-222.

Methé BA, Nelson KE, Deming JW, Momen B, Melamud E, Zhang X *et al* (2005). The psychrophilic lifestyle as revealed by the genome sequence of Colwellia psychrerythraea 34H through genomic and

proteomic analyses. *Proceedings of the National Academy of Sciences of the United States of America* **102:** 10913-10918.

Meyer F, Paarmann D, D'Souza M, Olson R, Glass E, Kubal M *et al* (2008). The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* **9:** 386.

Mykytczuk NC, Foote SJ, Omelon CR, Southam G, Greer CW, Whyte LG (2013). Bacterial growth at− 15° C; molecular insights from the permafrost bacterium Planococcus halocryophilus Or1. *The ISME journal* **7:** 1211-1226.

Niederberger TD, Sohm JA, Tirindelli J, Gunderson T, Capone DG, Carpenter EJ *et al* (2012). Diverse and highly active diazotrophic assemblages inhabit ephemerally wetted soils of the Antarctic Dry Valleys. *FEMS microbiology ecology* **82:** 376-390.

Parks DH, Beiko RG (2010). Identifying biologically relevant differences between metagenomic communities. *Bioinformatics* **26:** 715-721.

Piotrowska-Niczyporuk A, Bajguz A (2014). The effect of natural and synthetic auxins on the growth, metabolite content and antioxidant response of green alga Chlorella vulgaris (Trebouxiophyceae). *Plant Growth Regul* **73:** 57-66.

Pointing SB, Chan Y, Lacap DC, Lau MCY, Jurgens JA, Farrell RL (2009). Highly specialized microbial diversity in hyper-arid polar desert. *Proceedings of the National Academy of Sciences* **106:** 19964-19969.

Pointing SB, Belnap J (2012). Microbial colonization and controls in dryland systems. *Nature Reviews Microbiology* **10:** 551-562.

Rivkina EM, Friedmann EI, McKay CP, Gilichinsky DA (2000). Metabolic activity of permafrost bacteria below the freezing point. *Appl Environ Microb* **66:** 3230-3233.

Ruibal C, Gueidan C, Selbmann L, Gorbushina A, Crous P, Groenewald J *et al* (2009). Phylogeny of rock-inhabiting fungi related to Dothideomycetes. *Studies in Mycology* **64:** 123-133-S127.

Selbmann L, De Hoog G, Mazzaglia A, Friedmann E, Onofri S (2005). Fungi at the edge of life: cryptoendolithic black fungi from Antarctic desert. *Stud Mycol* **51:** 1-32.

Selbmann L, Grube M, Onofri S, Isola D, Zucconi L (2013). Antarctic epilithic lichens as niches for black meristematic fungi. *Biology* **2:** 784-797.

773 Sherwood AR, Presting GG (2007). Universal primers amplify a 23S rDNA plastid marker in eukaryotic
774 alage and cyanobacteria *Journal of Phycology* **43:** 605-608.

775
776 Smith JJ, Tow LA, Stafford W, Cary C, Cowan DA (2006). Bacterial diversity in three different Antarctic
777 cold desert mineral soils. *Microb Ecol* **51:** 413-421.

778
779 Steven B, Pollard WH, Greer CW, Whyte LG (2008). Microbial diversity and activity through a
780 permafrost/ground ice core profile from the Canadian high Arctic. *Environmental Microbiology* **10:** 3388-
781 3403.

782
783 Stomeo F, Makhalanyane TP, Valverde A, Pointing SB, Stevens MI, Cary CS *et al* (2012). Abiotic factors
784 influence microbial diversity in permanently cold soil horizons of a maritime-associated Antarctic Dry
785 Valley. *FEMS microbiology ecology*.

786
787 Sun H, Nienow J, McKay C (2010). The antarctic cryptoendolithic microbial ecosystem. *Life in Antarctic*
788 *deserts and other cold dry environments—astrobiological analogs Cambridge University Press,*
789 *Cambridge***:** 110-138.

790
791 Swanson MM, Reavy B, Makarova KS, Cock PJ, Hopkins DW, Torrance L *et al* (2012). Novel
792 bacteriophages containing a genome of another bacteriophage within their genomes. *PLoS One* **7:**
793 e40683.

794
795 Tamppari L, Anderson R, Archer P, Douglas S, Kounaves S, Mckay C *et al* (2012). Effects of extreme cold
796 and aridity on soils and habitability: McMurdo Dry Valleys as an analogue for the Mars Phoenix landing
797 site. *Antarctic Science* **24:** 211-228.

798
799 Tuorto SJ, Darias P, McGuinness LR, Panikov N, Zhang T, Häggblom MM *et al* (2014). Bacterial genome
800 replication at subzero temperatures in permafrost. *The ISME journal* **8:** 139-149.

801
802 Webb KM, DiRuggiero J (2013). Radiation resistance in extremophiles: fending off multiple attacks.
803 *Polyextremophiles*. Springer. pp 249-267.

804
805 Wei ST, Higgins CM, Adriaenssens EM, Cowan DA, Pointing SB Genetic signatures indicate widespread
806 antibiotic resistance and phage infection in microbial communities of the McMurdo Dry Valleys, East
807 Antarctica. *Polar Biology***:** 1-7.

808
809 Wierzchos J, de los Ríos A, Ascaso C (2013). Microorganisms in desert rocks: the edge of life on Earth.
810 *International Microbiology* **15:** 172-182.

811
812 Yergeau E, Hogues H, Whyte LG, Greer CW (2010). The functional potential of high Arctic permafrost
813 revealed by metagenomic sequencing, qPCR and microarray analyses. *The ISME journal* **4:** 1206-1214.

Yilmaz S, Allgaier M, Hugenholtz P (2010). Multiple displacement amplification compromises quantitative analysis of metagenomes. *Nat Meth* **7:** 943-944.

Zablocki O, van Zyl L, Adriaenssens EM, Rubagotti E, Tuffin M, Cary SC *et al* (2014). High-level diversity of tailed Phages, eukaryote-associated viruses, and virophage-like elements in the metaviromes of Antarctic soils. *Appl Environ Microb* **80:** 6888-6897.

Figure and Table Legends

Figure 1. University Valley permafrost and cryptoendolith niches

A. University Valley; B. University Valley cryptoendolith colonized zone, with adjacent mm ruler; C-D. Field cryptoendolith photos C. An exposed cryptoendolith in University Valley after a snowfall event D. Cryptoendolith community underneath the rocks surface utilizing low sunlight reflected by sandstone wall

Figure 2. Heterotrophic and photosystem activity at sub-zero temperatures

A. ^{14}C acetate mineralization detectable in cryptoendolithic at all temperatures tested, and in permafrost samples only above 5ºC. B-C. *Diplosphaera* (B) and *Stichococcus* (C) isolates demonstrating chloroplast autofluorescence activity after 200 days culturing at −5ºC.

Figure 3. Functional profiles of University Valley cryptoendoliths and permafrost

Level 1 SEED subsystems found in the cryptoendolith and permafrost metagenomes. Indicated by * are subsystems which were found to be biologically important, as inferred from statistical probability modelling using Statistical Analysis of Metagenomic (STAMP) v 2.0.8 (Parks and Beiko 2010) using *P*-value <0.05, for ratio of proportions (RP) effect size <2.00 and difference of proportions effect size <2.5.

Figure 4. Ordination of functional community composition in University Valley

Principle coordinate analyses using Bray-Curtis distances of the relative abundance of level 2 SEED subsystems in the University Valley permafrost (Upper) and cryptoendolith (Lower)

compared with other publically available metagenomes. Metagenomes used for comparison are outlines in Table S3. Arrows represent the relative abundance of level 2 subsystems.

Figure 5. Predicted functions in University Valley permafrost and cryptoendolithic systems

Table 1. University Valley permafrost and cryptoendolith metagenome statistics

Table 2. Abundant phyla and classes in University Valley metagenomes

Only Phyla which represent >1% of total reads are presented here.

Table 3. Key Nutrient Cycling gene(s)

No. of reads of key genes based on 60% protein identity, an e-value cut-off of e-5 and a minimum alignment length of 15 aas against the GenBank database.

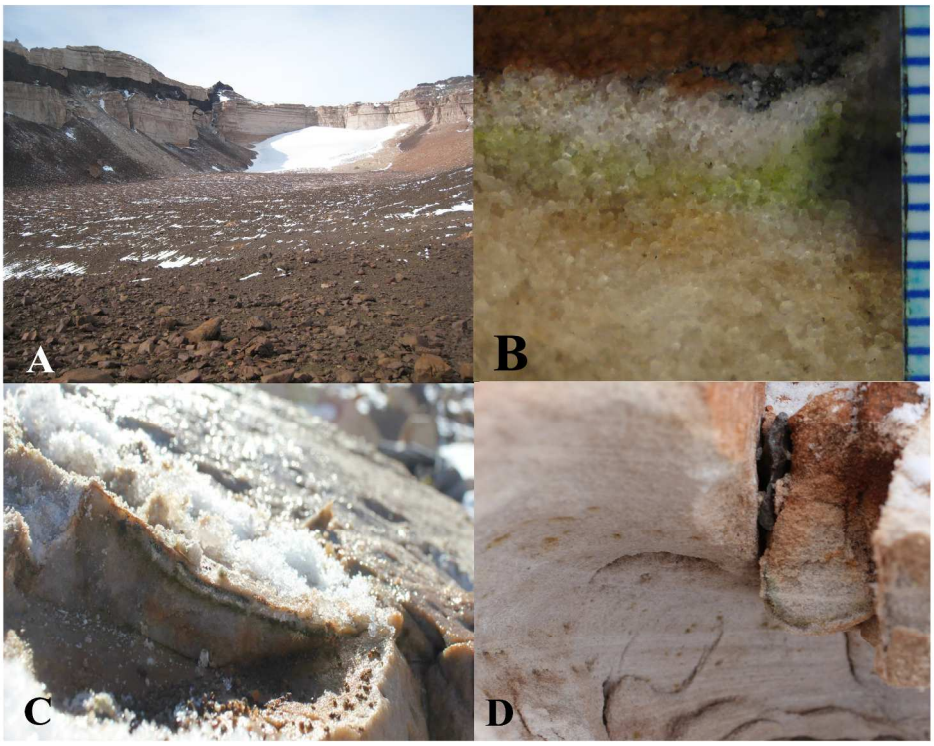

Figure 1. University Valley permafrost and cryptoendolith niches
A. University Valley; B. University Valley cryptoendolith colonized zone, with adjacent mm ruler; C-D. Field cryptoendolith photos C. An exposed cryptoendolith in University Valley after a snowfall event D. Cryptoendolith community underneath the rocks surface utilizing low sunlight reflected by sandstone wall

282x211mm (300 x 300 DPI)

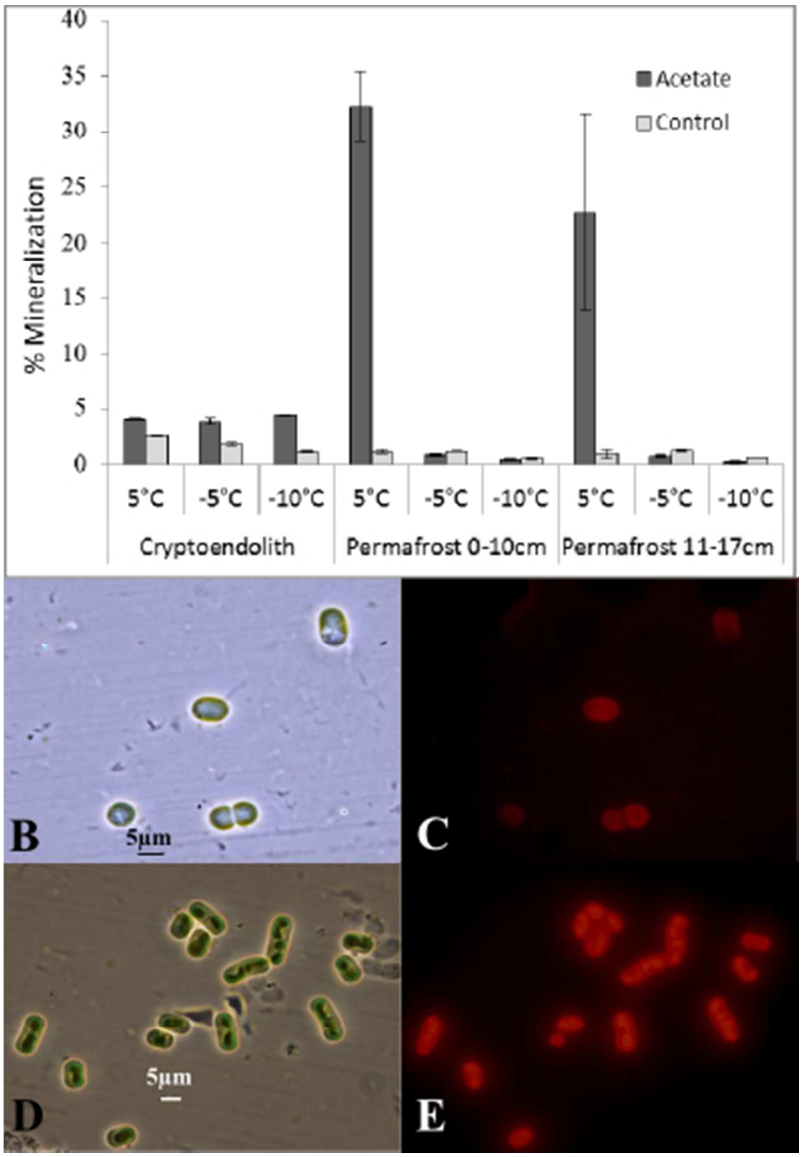

Figure 2. Heterotrophic and photosystem activity at sub-zero temperatures
A. 14C acetate mineralization detectable in cryptoendolithic at all temperatures tested, and in permafrost samples only above 5ºC. B-C. Diplosphaera (B) and Stichococcus (C) isolates demonstrating chloroplast autofluorescence activity after 200 days culturing at −5ºC.

105x151mm (96 x 96 DPI)

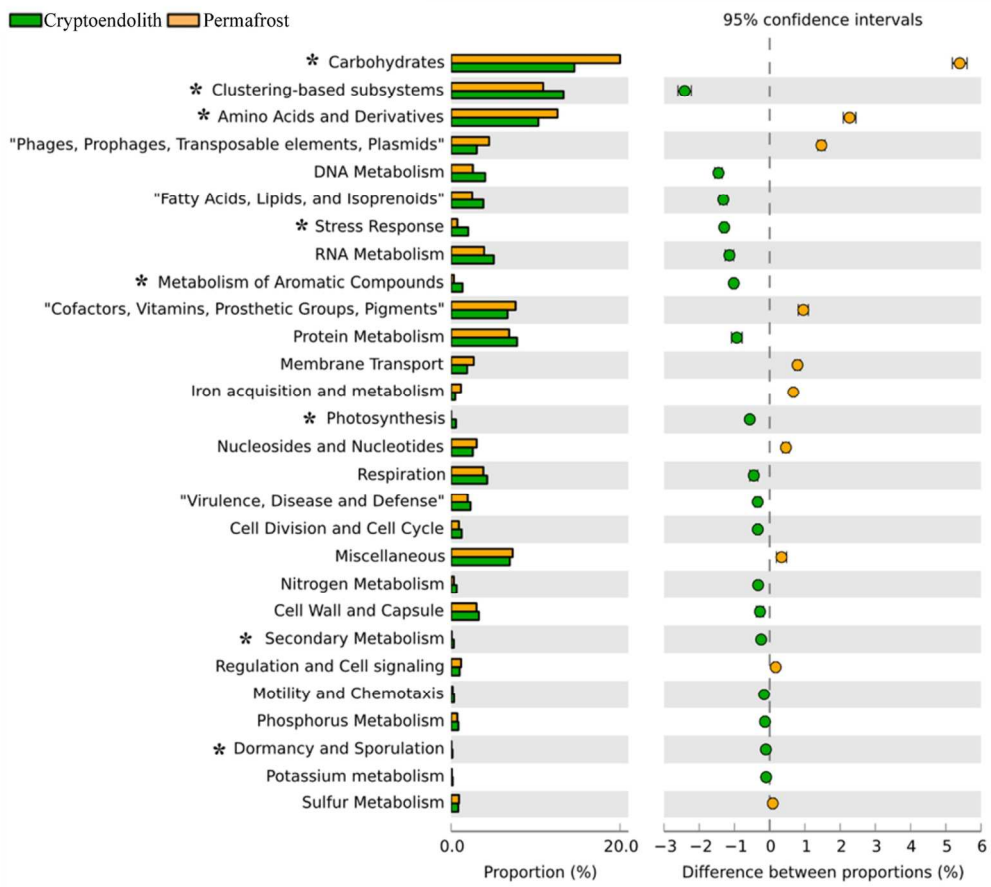

Figure 3. Functional profiles of University Valley cryptoendoliths and permafrost
Level 1 SEED subsystems found in the cryptoendolith and permafrost metagenomes. Indicated by * are subsystems which were found to be biologically important, as inferred from statistical probability modelling using Statistical Analysis of Metagenomic (STAMP) v 2.0.8 (Parks and Beiko 2010) using P-value <0.05, for ratio of proportions (RP) effect size <2.00 and difference of proportions effect size <2.5.

149x133mm (300 x 300 DPI)

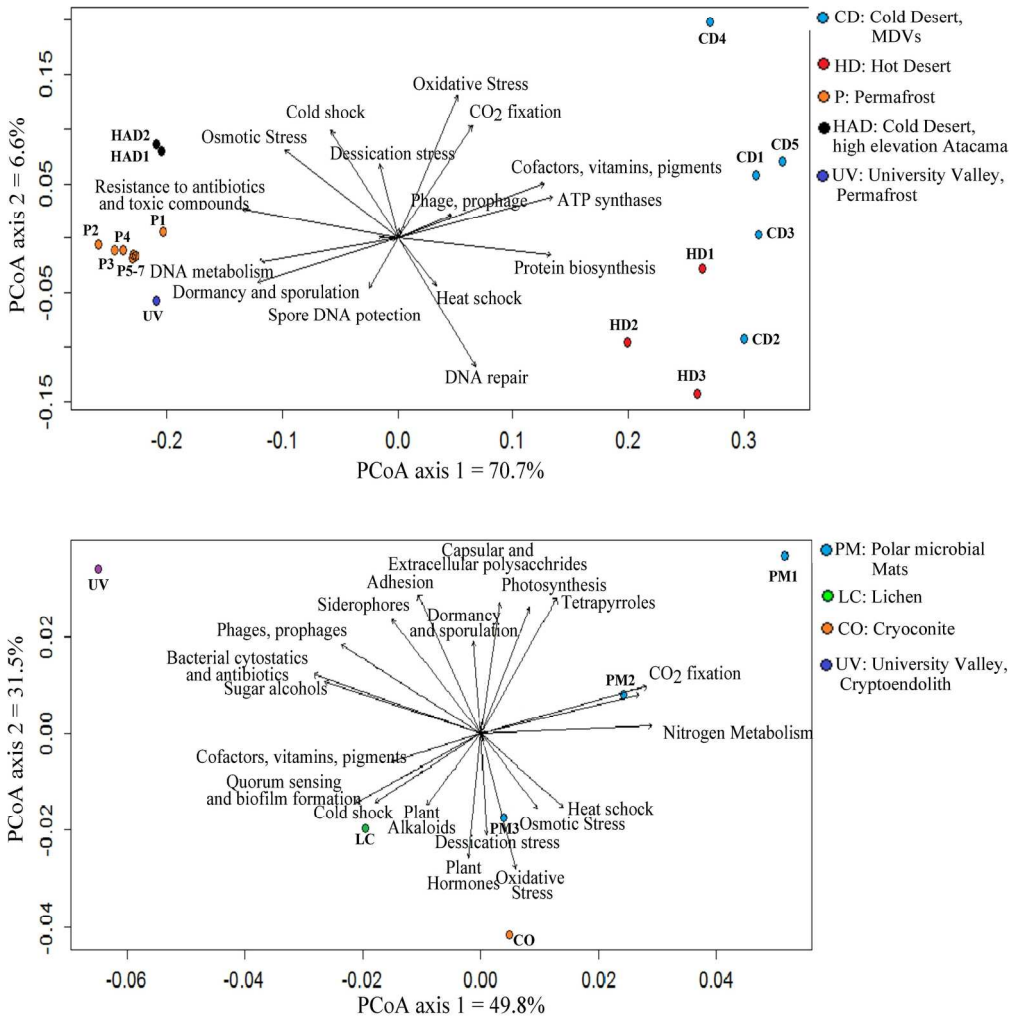

Figure 4. Ordination of functional community composition in University Valley
Principle coordinate analyses using Bray-Curtis distances of the relative abundance of level 2 SEED subsystems in the University Valley permafrost (Upper) and cryptoendolith (Lower) compared with other publically available metagenomes. Metagenomes used for comparison are outlines in Table S3. Arrows represent the relative abundance of level 2 subsystems. \r\n
184x202mm (300 x 300 DPI)

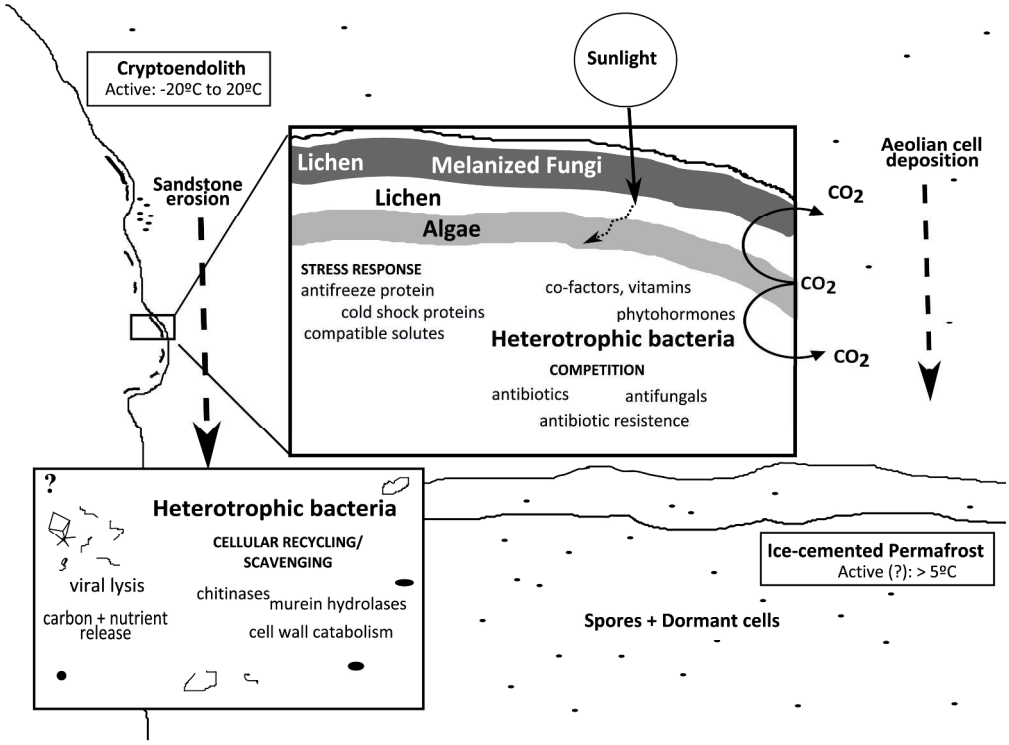

Figure 5. Predicted functions in University Valley permafrost and cryptoendolithic systems
247x183mm (300 x 300 DPI)

Table 1. University Valley permafrost and cryptoendolith metagenome statistics

Parameter	Cryptoendolith Metagenome	Permafrost Metagenome
Total no. of sequences before QC	1,293,156	3,124,825
No. of sequences that passed QC	1,112,128	737,531
Total sequence size (bp) after QC	309,810,374 bp	211,084,258 bp
Av. sequence length (bp) after QC	278 bp	286 bp
No. of predicted/identified protein features	770,392/ 193,269	146,715/ 63,452
No. of predicted/identified rRNA features	7,444/ 441	41,691/ 242
No. of identified functional categories	128,195	42,837
GC content (%)	50 %	54 %

Table 2. Abundant phyla and classes in University Valley metagenomes

Domain	Phylum [class]	Percent (%) abundance cryptoendolith metagenome	Percent (%) abundance permafrost metagenome
Eukaryota		**53.8**	**12.2**
	Ascomycota	45.0	9.0
	[Eurotiomycetes]	24.1	2.8
	[Sordariomycetes]	10.0	1.7
	[Dothideomycetes]	5.2	3.9
	[Leotiomycetes]	4.1	0.3
	Chlorophyta	3.4	<0.1
	[Trebouxiophyceae]	1.4	0
	[Chlorophyceae]	1.2	<0.1
	Basidiomycota	2.2	0.5
	[Tremellomycetes]	1.7	<0.1
	Streptophyta	2.0	0.1
Bacteria		**45.0**	**85.6**
	Actinobacteria	19.8	63.7
	(order) Actinomycetales	17.4	63.2
	Proteobacteria	10.3	13.6
	[Alphaproteobacteria]	4.6	5.2
	[Gammaproteobacteria]	2.0	2.8
	[Deltaproteobacteria]	1.9	0.4
	[Betaproteobacteria]	1.6	5.0
	Bacteroidetes	3.9	1.6
	[Sphingobacteria]	1.3	0.4
	[Cytophagia]	1.2	0.2
	Firmicutes	2.2	5.9
	[Clostridia]	1.2	0.8
	[Bacilli]	1.0	5.0
	Chloroflexi	2.0	0.0
	Acidobacteria	1.9	0.1
	Cyanobacteria	1.7	0.3
	(order) Chroococcales	0.8	0.2
	(order) Nostocales	0.4	<0.1
	(order) Oscillatoriales	0.2	<0.1
	(order) Gloeobacterales	0.2	<0.1
Viruses		**0.5**	**2.0**
Archaea		**0.4**	**<0.1**

Only Phyla which represent >1% of total reads are presented here.

Table 3. Key Nutrient Cycling gene(s)

	Pathway	Key gene(s)	Cryptoendolith No. of reads	Permafrost No. of reads
Carbon-Autotrophy	**Calvin-Benson Cycle**	ribulose-1,5-bisphosphate carboxylase/oxygenase phosphoribulokinase	127 14	0 0
	Reductive TCA Cycle	2-oxoglutarate:ferredoxin oxidoreductase ATP citrate lyase (aclB)	0 7	0 0
	Reductive acetyl-coA pathway	CO dehydrogenase/acetyl-CoA synthase (CO-DH)	0	0
	Hydroxypropionate cycle	acetyl-CoA/propionyl-CoA carboxylase (pcc) malonyl coA reductase	0 0	0 0
Carbon	**Methane Oxidation**	Methane monooxygenase (mmoX) Particulate methane monoxygenase (pmoA)	13 0	0 0
	Methanogenesis	Methyl coenzyme M reductase (mcrA)	0	0
	Acetogenesis	Formyltetrahydrofolate synthetase (FTHFS)	0	1
	Carbon monoxide	CO dehydrogenase/acetyl-CoA synthase CO-DH	0	0
	Glyoxalate pathway	isocitrate lyase malate synthase	29 63	326 78
Nitrogen	**Nitrogen Fixation**	Nitrogenase Reductase (nifH)	*0*	0
	Nitrification	Ammonia monooxygenase (amoA)	3	0
	Denitrification	Nitrate Reductase (narG, nasA, napA) Nitrite reductase (nirK, nirS, nirA, nirB, nrfA) Nitric Oxide reductase (norB, norVW) Nitrous Oxide reductase (NosZ)	66 62 0 0	62 56 0 0
	Mineralization	Glutamate dehydrogenase (gdh) Urea amidohydrolase (ureC)	120 1	89 0
Phosphorus	**Phosphate metabolism**	Alkaline Phosphatases (phoA and PhoX) Phosphate-specific transport (Pst operon)	81 103	3 861
	Phosphonate metabolism	Phosphonoacetaldehyde hydrolase (phnX)	0	1
	Polyphosphonate metabolism	(polyphosphatase kinase (ppK) Exopolyphosphatase (ppX)	0 27	0 7

No. of reads of key genes based on 60% protein identity, an e-value cut-off of e-5 and a minimum alignment length of 15 aas against the GenBank database